PCCP

# Multistructural Microiteration Combined with QM/MM-ONIOM Electrostatic Embedding



| | |
|---|---|
| Journal: | *Physical Chemistry Chemical Physics* |
| Manuscript ID | CP-ART-05-2022-002270 |
| Article Type: | Paper |
| Date Submitted by the Author: | 19-May-2022 |
| Complete List of Authors: | Suzuki, Kimichi; Hokkaido University,<br>Maeda, Satoshi; Faculty of Science, Hokkaido University, Chemistry |
| | |

# Multistructural Microiteration Combined with QM/MM-ONIOM Electrostatic Embedding

Kimichi Suzuki*[a,b,c] and Satoshi Maeda*[a,b,c,d]

[a] *Institute for Chemical Reaction Design and Discovery* (WPI-ICReDD)*, Hokkaido University, Sapporo 001-0021, Japan*

[b] *Department of Chemistry, Faculty of Science, Hokkaido University, Sapporo 060-0810, Japan*

[c] *JST, ERATO Maeda Artificial Intelligence for Chemical Reaction Design and Discovery Project, Sapporo 060-0810, Japan*

[d] *Research and Services Division of Materials Data and Integrated System (MaDIS), National Institute for Materials Science (NIMS), Tsukuba 305-0044, Japan*

**Abstract:** Multistructural microiteration (MSM) is a method to take account of contributions of multiple surrounding structures in a geometrical optimization or reaction path calculation using the quantum mechanics/molecular mechanics (QM/MM) ONIOM method. In this study, we combined MSM with the electrostatic embedding (EE) scheme of the QM/MM-ONIOM method by extending its original formulation for mechanical embedding (ME). MSM-EE takes account of the polarization in the QM region induced by point charges assigned to atoms in the multiple surrounding structures, where the point charges are scaled by the weight factor of each surrounding structure determined through MSM. The performance of MSM-EE was compared with that of the other methods, i.e., ONIOM-ME, ONIOM-EE, and MSM-ME, by applying them to three chemical processes: (1) chorismate-to-prephenate transformation in aqueous solution, (2) the same transformation as (1) in

an enzyme, and (3) hydroxylation in *p*-hydroxybenzoate hydroxylase. These numerical tests of MSM-EE yielded barriers and reaction energies close to experimental values with computational costs comparable to those of the other three methods.

## 1. Introduction

Recent advances in computer hardware and quantum chemical calculation software have made it relatively easy to theoretically elucidate mechanisms of organic reactions.[1-3] In many such studies, transition states (TSs) for ten or more elementary processes have been computed to draw reliable conclusions. Moreover, there have been cases where hundreds or more elementary steps were examined to construct a reaction path network and discuss detailed reaction kinetics.[4-6] However, it is still not easy to conduct a similar theoretical study of enzymes taking account of large-scale changes in the surrounding protein structure.

The quantum mechanics (QM)/molecular mechanics (MM) method has been widely used in studies on large molecular systems. This approach efficiently computes the energy of a large system by partitioning the entire system into a reaction center and another part and applying different levels of theory to each.[7-15] QM/MM methods can be categorized into additive[7-10] and subtractive[11-12] types. The generally known QM/MM methods are based on the additive scheme. On the other hand, one popular QM/MM method of the latter type is our own *N*-layered integrated molecular orbital molecular mechanics (ONIOM) method.[11,15] ONIOM is general and flexible because it can divide entire system into two or more layers like an onion and apply different QM and/or MM methods to each region. QM/MM methods are now available in many *ab initio* quantum chemistry calculation programs.[16-21] Moreover, many *ab initio* quantum chemistry calculation programs[25-32] and molecular dynamics (MD) simulation programs[25-32] have an interface that performs a QM/MM calculation in cooperation with the other program.

There are two major components in the QM/MM methods: the scheme of incorporating the electrostatic interaction between the QM and MM parts and the technique to describe the structure of

the MM part. Several methods have been proposed for them. The more advanced scheme for incorporating the electrostatic interaction between the QM and MM parts is polarizable embedding.[33, 34] This scheme takes account of mutual polarization between the MM part and the QM part. In order to introduce the polarization effect induced by QM part, three basic kinds of polarizable model have been proposed: fluctuating charge,[35] Drude oscillator, [36] and induced dipole models.[37] In addition, flexible embedding, which allows partial charge transfer between QM and MM parts, has been also proposed. [38]

In many studies, on the other hand, mechanical embedding (ME)[11] or electrostatic embedding (EE)[10, 39,40] has been adopted to conveniently incorporate the electrostatic interaction between the QM and MM parts. The ME scheme evaluates the electrostatic interaction with the MM method. In the EE scheme, the electrostatic interaction is computed with the QM method by including partial charges on surrounding atoms in the QM Hamiltonian. EE often provides more reasonable chemical descriptions than ME.[40,41] Therefore, EE has been widely used to study reaction mechanisms.

Geometrical microiteration has been widely used to efficiently optimize a whole system structure.[43,44] Before taking an optimization step for QM atomic coordinates, microiteration fully optimizes positions of atoms in the MM part under the condition that the QM atomic coordinates are fixed. The merit of microiteration is that it greatly reduces the number of QM calculations required before to reach convergence.[45] Thus, this technique has been used not only in geometrical optimizations but also in reaction path calculations.[14,15] Furthermore, automated reaction path searches have been applied to large molecular systems by combining QM/MM-ONIOM and geometrical microiteration.[45,46] With microiteration, geometrical optimization and reaction path calculations can be done routinely even for enzyme reactions.

However, microiteration often encounters difficulty in choosing the surrounding MM structure. Usually, a single surrounding structure prepared on the basis of a crystal structure and/or

an MD simulation is used. Difficulty can arise when the volume of the QM region drastically changes along a reaction coordinate or when a large-scale surrounding structural change occurs during a chemical reaction.[47-50] To solve this problem, we recently proposed a technique called multistructural microiteration (MSM). MSM describes the structure of the MM part as the average of multiple surrounding structures and takes account of their contributions during geometrical optimization and reaction path calculations.[51,52] This technique is just a simple extension of microiteration, and the computational cost is comparable to that of conventional microiteration. Despite the simplicity, MSM has been successful in describing large-scale surrounding structural transitions such as from open- to closed-loop conformation and vice versa during chemical reactions.[52]

MSM and EE are a reasonable combination for routinely conducting geometrical optimization and reaction path calculations. However, MSM has so far been based only on ME. In this study, we first develop the MSM-EE scheme by extending the original MSM-ME. The key idea is to combine MSM-ME and ONIOM-EE using partial charges scaled by the weighted sum of surrounding structures. Second, we demonstrate its performance by applying it to three chemical reactions: the transformation of chorismate into prephenate 1) in an aqueous solution and 2) in an enzyme, and 3) hydroxylation catalyzed by *p*-hydroxybenzoate hydroxylase. These reaction systems have been used for benchmark calculations of large molecular systems[53-59] because experimental kinetics data are available.[60-67] Barriers and reaction energies obtained through MSM-EE calculations are compared with experimental values and those obtained through ONIOM-ME, ONIOM-EE, and MSM-ME calculations.

## 2. Theory

In the original ONIOM method,[11,15] the total ONIOM-ME energy ($E^{\text{ONIOM-ME}}$) in the two-layer partition (QM and MM regions) is expressed as

$$E^{\text{ONIOM - ME}} = E^{\text{model - QM - ME}} - E^{\text{model - MM - ME}} + E^{\text{real - MM}}, \quad (1)$$

where $E^{\text{model-QM-ME}}$, $E^{\text{model-MM-ME}}$, and $E^{\text{real-MM-ME}}$ correspond to the QM energy of the model system, MM energy of the model system, and MM energy of the entire system, respectively. The QM Hamiltonian in the model system is given by

$$\hat{H}^{\text{model - QM - ME}} = -\frac{1}{2}\sum_{i}\nabla_i^2 + \sum_{ij}\frac{1}{r_{ij}} - \sum_{i\mu}\frac{Z_\mu}{r_{i\mu}} + \sum_{\mu\nu}\frac{Z_\mu Z_\nu}{R_{\mu\nu}}, \quad (2)$$

where $i$, $\mu$, $Z_\mu$, $r_{ij}$, and $R_{\mu\nu}$ refer to $i$th electron, $\mu$th nucleus, nuclear charge on the $\mu$th nucleus, distance between the $i$th and $j$th electrons, and distance between the $\mu$th and $\nu$th nuclei, respectively. For example, the energy functions for the AMBER force field [25] model system ($E^{\text{model-MM-ME}}$) and for the entire system ($E^{\text{real-MM}}$) in the ME scheme are

$$E^{\text{model - MM - ME}} = \sum_{\text{bonds}} k_r(r - r_{\text{eq}})^2 + \sum_{\text{angles}} k_\theta(\theta - \theta_{\text{eq}}) + \sum_{\text{dihedral}} \frac{V_n}{2}[1 + \cos(n\phi - \xi)]$$

$$+ \sum_{\mu<\nu}\left[\left(\frac{A_{\mu\nu}}{R_{\mu\nu}^{12}} - \frac{B_{\mu\nu}}{R_{\mu\nu}^{6}}\right) + \frac{q_\mu q_\nu}{\epsilon R_{\mu\nu}}\right] \quad (3)$$

and

$$E^{\text{real - MM}} = \sum_{\text{bonds}} k_r(r - r_{\text{eq}})^2 + \sum_{\text{angles}} k_\theta(\theta - \theta_{\text{eq}}) + \sum_{\text{dihedral}} \frac{V_n}{2}[1 + \cos(n\phi - \xi)]$$

$$+ \sum_{I<J}\left[\left(\frac{A_{IJ}}{R_{IJ}^{12}} - \frac{B_{IJ}}{R_{IJ}^{6}}\right) + \frac{q_I q_J}{\epsilon R_{IJ}}\right], \quad (4)$$

where $i$, $\mu$, $Z_\mu$, $r_{ij}$, and $R_{\mu\nu}$ refer to $i$th electron, $\mu$th nucleus, nuclear charge on the $\mu$th nucleus, distance between the $i$th and $j$th electrons, and distance between the $\mu$th and $\nu$th nuclei, respectively. The first three terms in eqs. (3) and (4) indicate the classical force fields for bond stretching, bond bending, and dihedral motions, respectively, while the last two terms represent non-covalent bond terms including Lennard–Jones and Coulomb interactions, respectively.

In the electrostatic embedding (EE) scheme,[39,40] the ONIOM-EE energy expression is

$$E^{\text{ONIOM - EE}} = E^{\text{model - QM - EE}} - E^{\text{model - MM - EE}} + E^{\text{real - MM}}, (5)$$

where

$$\hat{H}^{\text{model - QM - EE}} = \hat{H}^{\text{model - QM - ME}} - \sum_i \sum_I \frac{q_I}{r_{iI}} + \sum_\mu \sum_I \frac{Z_\mu q_I}{R_{\mu I}}, (6)$$

and

$$E^{\text{model - MM - EE}} = E^{\text{model - MM - ME}} + \sum_\mu \sum_I \frac{q_\mu q_I}{R_{\mu I}}. (7)$$

In ONIOM-EE, the electrostatic effect between QM and MM atoms is described by the second term in eq. (6), while the second term in eq. (7) is used to cancel out the electrostatic interaction computed with the lower level of theory. In MSM-ME[51], $E^{\text{MSM-ME}}$ is expressed as

$$E^{\text{MSM - ME}} = E^{\text{model - QM - ME}} - E^{\text{model - MM - ME}} + \sum_s^N \omega^{(s)} E^{\text{real - MM}(s)}, (8)$$

where $N$ is the number of surrounding structures considered, and $E^{\text{real-MM}(s)}$ corresponds to the MM energy of the $s$th surrounding structure. The weight $\omega^{(s)}$ is given by

$$\omega^{(s)} = \frac{e^{-\beta E^{\text{real} - \text{MM}(s)}}}{\sum_t^N e^{-\beta E^{\text{real} - \text{MM}(t)}}}, (9)$$

where $\beta = 1/k_B T^{\text{MSM}}$, in which $k_B$ is the Boltzmann constant and $T^{\text{MSM}}$ is a model temperature parameter. The concept of MSM is that reaction-center atoms are treated in the mean field generated by the weighted sum of the surrounding structures, while the surrounding part is described as the weighted sum of multiple surrounding structures. In the previously proposed MSM method, electrostatic interaction between QM and MM regions is computed with the lower level of theory and ONIOM-ME in eq. (1). Thus, the electron density in the QM region is not polarized by surrounding atoms. To take account of polarization induced in the QM region by multiple surrounding atoms, our idea is to combine eqs. (6) and (8) as follows:

$$E^{\text{MSM - EE}} = E^{\text{MSM - model - QM - EE}} - E^{\text{MSM - model - MM - EE}} + \sum_{s}^{N} \omega^{(s)} E^{\text{real - MM}(s)}, \quad (10)$$

where

$$\hat{H}^{\text{MSM - model - QM - EE}} = \hat{H}^{\text{model - QM - ME}} - \sum_{s}^{N} \omega^{(s)} \sum_{i} \sum_{J} \frac{q_j^{(s)}}{r_{iJ^{(s)}}} + \sum_{s}^{N} \omega^{(s)} \sum_{I} \sum_{J} \frac{Z_I q_j^{(s)}}{R_{IJ^{(s)}}}, \quad (11)$$

and

$$E^{\text{MSM - model - MM - EE}} = E^{\text{model - MM - ME}} + \sum_{s}^{N} \omega^{(s)} \sum_{I} \sum_{J} \frac{q_I q_j^{(s)}}{R_{IJ^{(s)}}}. \quad (12)$$

In our approach, partial charges scaled by the *s*th weight computed in eq. (9) are used in the second and third terms in eq. (11) and the second term in eq. (12). If a single surrounding structure is used or only a single surrounding structure dominates reaction path searches, eqs. (11) and (12) are identical to the ONIOM-EE scheme. Using scaled partial charges efficiently includes electrostatic interaction in the wavefunction without performing QM calculations *N* times.

## 3. Computational details

In this study, three test calculations were performed: Claisen rearrangement of chorismate to prephenate in aqueous solution and in the *Bacillus subtilis* chorismite mutase[62], hydroxylation catalyzed by *p*-hydroxybenzoate hydroxylase[67] as shown in Figure 1.

### 3.1. Setup for Claisen rearrangement of chorismate to prephenate in aqueous solution and in the *Bacillus subtilis* chorismite mutase

For the reaction in aqueous solution, a cubic box with the length 40 Å, which contains a chorismate structure optimized in the gas phase, two Na ions as the counterions, and 7488 environment water molecules, was prepared. For the enzymatic reaction, the protein structure was constructed starting from the X-ray crystal structure (PDB code: 2CHT).[62] Hydrogen atoms were added using the program PROPKA.[68,69] Then, a cubic box with the length 80 Å, which contains the

enzyme structure, 10 Na ions, and 15722 environment water molecules, was prepared. Following preliminary preparations, replica-exchange molecular dynamics (MD) simulations with fixed QM coordinates were performed for 20 replicas spaced over the temperature range 300 – 400 K with 10,000,000 steps, where the QM atoms consisted of substrate (chorismate) both in aqueous solution and in the enzyme. After the NPT MD equilibration simulation, the model consisting of the chorismate structure, 2277 environment water molecules existing within 8.0 Å of the chorismate, and 2 Na ions was employed. For the enzymatic reaction, after the NPT MD equilibration simulation, the model consisting of the enzyme structure, 2789 environment water molecules existing within 5.0 Å of the enzyme, and 10 Na ions was employed. The five lowest MM energy configurations for initial surrounding structures were chosen from the geometric microiteation scheme for MSM calculations, while the surrounding structure of the lowest MM energy was used as the initial surrounding structure in the conventional ONIOM calculations. The replica-exchange MD simulations were performed with the weak coupling thermostat by Berendsen, while NPT simulations were performed with the Berendsen barostat and the Andersen thermostat. During the MD simulations, electrostatic interactions were evaluated by the Particle Mesh Ewald method setting the cutoff distance as 8.0 Å for van der Walls (vdW) interactions. All MD simulations were performed under periodic boundary conditions (PBCs) with $dt$ = 1.0 (fs) in the fully solvated cubic cell water box in the TIP3P model using program suites of AMBER version 12.[25] PBCs were used only in the MD simulations to obtain snapshots of initial surrounding structures. Hereafter, the five prepared surrounding structures of both aqueous solutions and enzymes are labeled as "$s$". The relative energies of the five surrounding structures for Claisen rearrangement were 0.0 kcal mol$^{-1}$ ($s$=1), 1.7 kcal mol$^{-1}$ ($s$=2), 4.0 kcal mol$^{-1}$ ($s$=3), 5.1 kcal mol$^{-1}$ ($s$=4), and 6.1 kcal mol$^{-1}$ ($s$=5) for aqueous solutions and 0.0 kcal mol$^{-1}$ ($s$=1), 2.6 kcal mol$^{-1}$ ($s$=2), 3.1 kcal mol$^{-1}$ ($s$=3), 3.5 kcal mol$^{-1}$ ($s$=4), and 7.4 kcal mol$^{-1}$ ($s$=5) for enzymes.

### 3.2. Setup for hydroxylation by *p*-hydroxybenzoate hydroxylase

The setup procedures for *p*-hydroxybenzoate hydroxylase were essentially the same as for Claisen rearrangement. The protein structure was taken from X-ray crystal structure (PDB code: 1PBE).[67] After adding hydrogen atoms to the enzyme, the cubic box with the length 90 Å, which contains an enzyme structure, 3 Na ions, and 37502 environment water molecules, was prepared. Afterward, replica-exchange and equilibrated NPT MD simulations were performed, and the enzyme structure, environment water molecules (2838) within 5.0 Å of the enzyme, and 3 Na ions were employed as the enzyme model structure. The five lowest MM energy configurations for initial surrounding structures were chosen from the geometric microiteation scheme for MSM calculations. The relative energies of the five surrounding structures were 0.0 kcal mol$^{-1}$ ($s$=1), 5.6 kcal mol$^{-1}$ ($s$=2), 16.1 kcal mol$^{-1}$ ($s$=3), 20.4 kcal mol$^{-1}$ ($s$=4), and 22.9 kcal mol$^{-1}$ ($s$=5).

All QM calculations were performed with the program ORCA 4.2.0[21], while all MM calculations were performed with the program Gaussian09.[16] Both ONIOM and MSM with ME and EE were implemented in the developer version of the program GRRM.[70,71] Although these methods in GRRM are able to perform automated searches of multiple reaction pathways, we used a step-by-step procedure with a targeted path search to compare their performances. First, an artificial force induced reaction (AFIR) method[72] was used to obtain rough initial reaction paths. The artificial force was applied between two fragments (denoted as circles in Figure 1). The model collision parameter in AFIR was set to 200 kJ mol$^{-1}$. $T^{\mathrm{MSM}}$ in MSM calculations was set as 20,000 K for Claisen rearrangement and 50,000 K for hydroxylation. To take into account the contributions of all surrounding structures during AFIR calculations, we determined the large $T^{\mathrm{MSM}}$ value at initial structures stage so that the weight factor of each surrounding structures satisfied the following conditions[51]:

$$\omega^{(s)} = \frac{e^{-\beta E^{\mathrm{real-MM}(s)}}}{\sum_t^N e^{-\beta E^{\mathrm{real-MM}(t)}}} \geq \frac{1}{N} \times 0.9 \quad (13)$$

Second, locally update plane (LUP) path optimization[73] was used to relax the initial reaction path obtained in the first step. During LUP path optimization, all energy maxima were directly optimized for the actual TS using a quasi-Newton algorithm. Finally, IRC path optimization was performed. The enthalpy and free energy were computed using harmonic approximation. QM and MM calculations were performed using B3LYP/6-31+G(d,p) with the RIJCOSX approximation[74,75] and AMBER force field[25], respectively. In all schemes, the vdW interaction between QM and MM regions is classically evaluated with the weight factor determined by eq. (9), while electrostatic interaction between QM and MM regions is treated by quantum mechanically in ONIOM-EE and MSM-EE schemes. To handle a QM and MM boundary region in the cofactor flavin hydroperoxide in *p*-hydroxybenzoate hydroxylase, the link atom method is employed in this study.[11] In order to avoid over polarization of the electron density in the QM part, the charge redistribution approach[76] is applied to the QM and MM boundary region as described in detail in S1. The atomic charges by a restrained electrostatic potential in QM regions for MM calculations were evaluated at optimized reactant structure. MM parameters were taken from generalized amber force field files, while missing parameters were created using Seminario's scheme.[77]

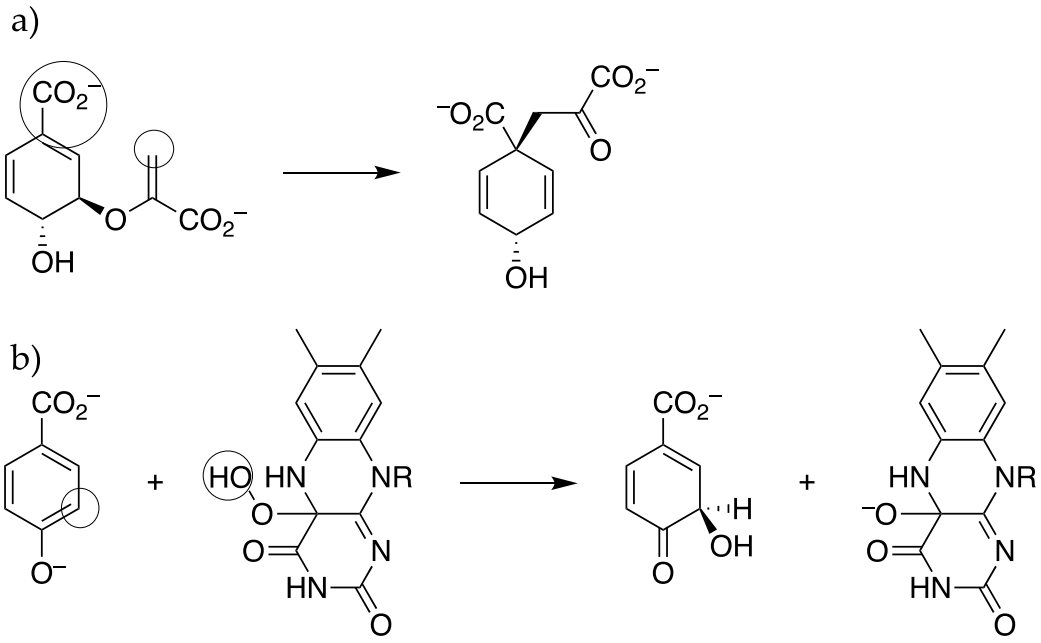

**Figure 1**. (a) Schematic of Claisen rearrangement in aqueous solution and in the *Bacillus subtilis* chorismite mutase, and (b) hydroxylation in *p*-hydroxybenzoate hydroxylase.

## 4. Results and discussion

### 4.1 Claisen rearrangement in aqueous solution

Figures 2(a)–(c) show the relative energies obtained through all methods and weight factors ($\omega^{(s)}$) using MSM along an AFIR path. In Figure 2(a), the energy profiles from all methods are almost the same around the first 15 AFIR steps. After that, the AFIR paths in all methods were obtained as a single peak, where the peak relative energies decrease in the order ONIOM-ME (31.5 kcal mol$^{-1}$) < MSM-ME (33.1 kcal mol$^{-1}$) < MSM-EE (36.6 kcal mol$^{-1}$) < ONIOM-EE (41.8 kcal mol$^{-1}$). Around the peak in MSM-ME and MSM-EE, the favorable surrounding structures changed into $s$=5 for MSM-ME and $s$=4 as shown in Figures 2(b) and 2(c), respectively. Thus, the relative energy differences between ONIOM-ME and MSM-ME or between ONIOM-EE and MSM-EE are mainly due to the difference in favorable surrounding structure. The peak shows a clear difference between ME and EE. The approximate reaction energies are –32.0 kcal mol$^{-1}$ from ONIOM-ME, –41.2 kcal mol$^{-1}$ from MSM-ME, –7.6 kcal mol$^{-1}$ from ONIOM-EE, and–23.9 kcal mol$^{-1}$ from MSM-EE. This implies that the product state in ME is more stable than in EE.

Figures 3(a)–(c) show the IRC paths obtained by all methods and $\omega^{(s)}$ by MSM. TS structures were obtained through LUP path optimization. The optimized LUP paths and variation of $\omega^{(s)}$ are shown in S2, while geometric parameters at the stationary points are shown in S3. The computed barrier heights and reaction energies are listed in Table 1. The MSM-EE barrier is the highest in Figure 3(a), where the order of obtained barrier heights is ONIOM-ME (17.2 kcal mol$^{-1}$) ≈ MSM-ME (17.5 kcal mol$^{-1}$) ≈ ONIOM-EE (19.1 kcal mol$^{-1}$) < MSM-EE (24.4 kcal mol$^{-1}$). The

reaction energies from ONIOM- and MSM-ME were much lower than those from EE; those values were –30.2 kcal mol$^{-1}$ from ONIOM-ME, –34.6 kcal mol$^{-1}$ from MSM-ME, –15.1 kcal mol$^{-1}$ from ONIOM-EE, and –15.8 kcal mol$^{-1}$ from MSM-EE. In Table 1, the free energy, enthalpy, and reaction energy from MSM-EE agree reasonably well with the experimental results. However, both ME schemes overestimate/underestimate the reaction energy/barrier height compared with experimental results. While ONIOM-EE provides a reasonable reaction energy, the barrier height is underestimated by about 7 kcal mol$^{-1}$. Continuum solvation models that have been used for many organic reactions can efficiently evaluate solvent effects without considering a large number of explicit solvent molecules and their arrangements and thus are much more efficient than the ONIOM/MSM-ME/EE schemes. However, the conductor-like polarizable continuum model without explicit water molecules estimated the enthalpy barrier height and reaction enthalpy values as 14.9 kcal mol$^{-1}$ and -20.2 kcal mol$^{-1}$, respectively, underestimating their magnitudes. Thus, it is important to take into account the multiple surrounding structural contributions and electrostatic interactions during a reaction path search.

Regarding multiple surrounding structural contributions, the MSM methods provided lower absolute energies than the ONIOM methods throughout the IRC path. For example, the total energies for the reactant structure were –849.96898 (a.u.) for ONIOM-EE, –849.99787 (a.u.) for MSM-EE, –850.060474 (a.u.) for ONIOM-ME, and –850.06857 (a.u.) for MSM-ME. Total energy lowering by MSM treatments between same embedding schemes is because favorable surrounding structures changed during AFIR, LUP, and IRC path calculations, and therefore no longer included a contribution from $s$=1. While a single surrounding structure ($s$=4) dominated through the IRC path in MSM-ME as shown in Figure 3(b), two surrounding structures were indispensable in MSM-EE, because the most favorable surrounding structural change from $s$=5 to $s$=4 around the product state was observed on the IRC path as shown in Figure 3(c).

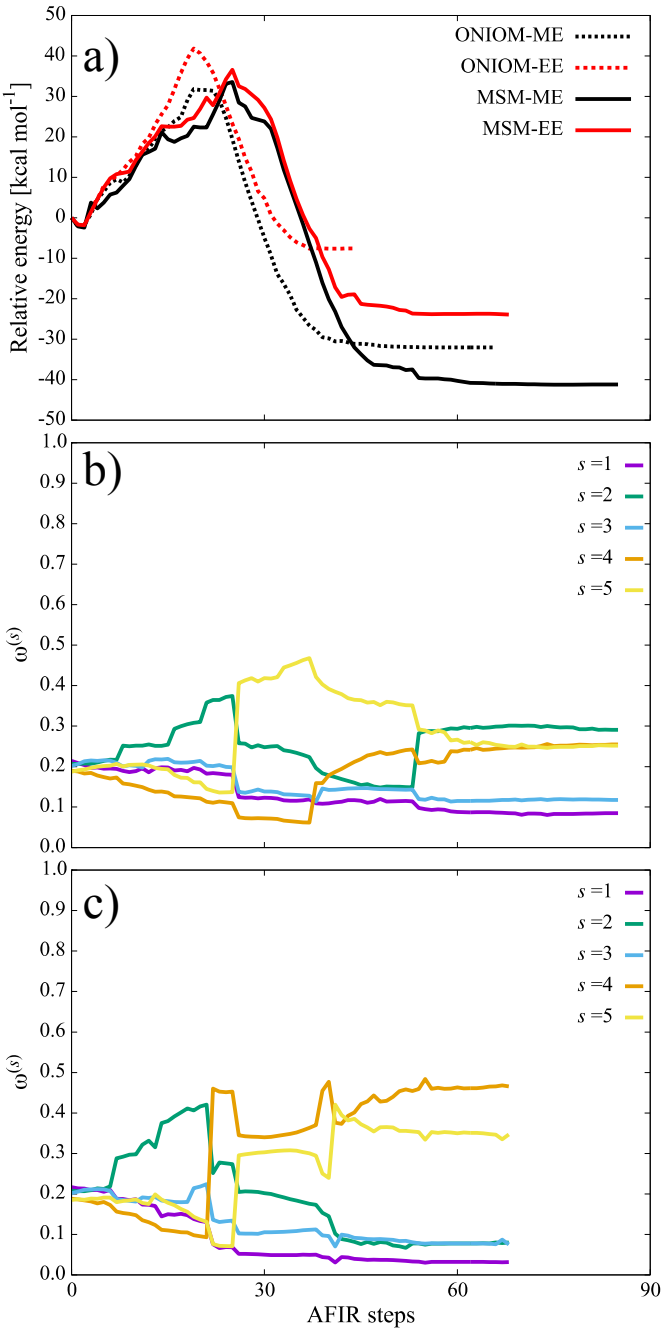


**Figure 2**. (a) Relative energies and variations of $\omega^{(s)}$ (b) in MSM-ME and (c) in MSM-EE along the AFIR path of Claisen rearrangement in aqueous solution.

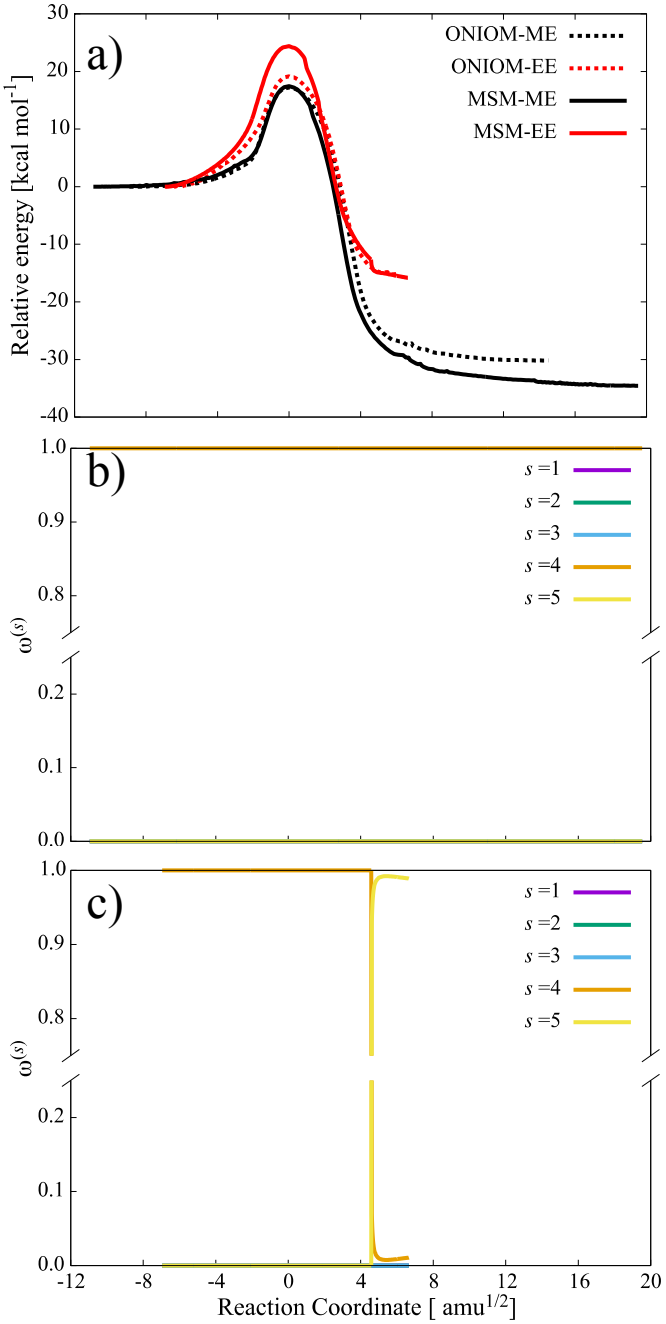


**Figure 3**. (a) Relative energies and variations of $\omega^{(s)}$ (b) in MSM-ME and (c) in MSM-EE along the IRC path of Claisen rearrangement in aqueous solution.

**Table 1**. Potential energy (*E*), free energy (*G*), and enthalpy (*H*) barrier heights, and computed and experimental reaction energies of Claisen rearrangement in aqueous solution. The units are kcal mol$^{-1}$.

| | Barrier height | | | Reaction energy | | |
|---|---|---|---|---|---|---|
| | *E* | *G* | *H* | *E* | *G* | *H* |
| ONIOM-ME | 17.2 | 15.7 | 15.3 | -30.2 | -30.8 | -30.4 |
| ONIOM-EE | 19.2 | 17.8 | 17.5 | -15.1 | -15.4 | -14.9 |
| MSM-ME | 17.5 | 15.5 | 15.5 | -34.6 | -34.9 | -34.9 |
| MSM-EE | 24.4 | 23.1 | 22.9 | -15.8 | -16.2 | -15.7 |
| Exp[60] | | 24.5 | 20.7 | | | -13.2 |

### 4.2 Claisen rearrangement in an enzyme

Figures 4(a)–(c) show the relative energies from all methods and $\omega^{(s)}$ from MSM methods along the AFIR path. The energy profiles throughout the AFIR paths are almost the same between ONIOM-ME and MSM-ME and between ONIOM-ME and MSM-EE. The relative energies of the highest peak decrease in the order ONIOM-ME (14.1 kcal mol$^{-1}$) < MSM-ME (15.2 kcal mol$^{-1}$) < ONIOM-EE (23.4 kcal mol$^{-1}$) < MSM-EE (25.4 kcal mol$^{-1}$). Clearly, the main difference between ME and EE is that ME tends to stabilize transition and product states. This tendency is the same as for aqueous solution. In Figures 4(b) and (c), the two surrounding structures *s*=1 and *s*=3 dominate the first 30 AFIR steps. After 30 steps, the surrounding structure *s*=1 maintains a large contribution in MSM-ME and MSM-EE, resulting in AFIR paths almost identical to those in ONIOM methods.

Figures 5(a)–(c) show IRC paths in all methods and □□s□□in the□MSM schemes. LUP path optimizations, which are shown in Figure S4, were performed before IRC calculations. The optimized geometric parameters at the stationary points are also shown in S5. Table 2 lists the barrier heights and reaction energies. The IRC paths are almost the same between ONIOM-ME and MSM-ME and between ONIOM-EE and MSM-EE in the range from –8 to 4 (Å amu $^{1/2}$) in reaction

coordinates. The surrounding structure *s*=3 in MSM-ME and MSM-EE dominates throughout the IRC paths. These results clearly reflect the results for AFIR paths. A small multi-surrounding structural contribution is observed around 2.5 (Å amu $^{1/2}$) in MSM-ME. However, the IRC profile is not largely affected by its contribution. In fact, the computed barrier heights are the same between ONIOM-ME and MSM-ME and between ONIOM-EE and MSM-EE within 1 kcal mol$^{-1}$. Because multi-surrounding structural contributions in MSM are small on IRC paths, differences in absolute total energy between ONIOM-ME and MSM-ME and between ONIOM-EE and MSM-EE are also rather small. The total energies in the reactant state are –888.46381 (a.u.) for ONIOM-EE, –888.46577 (a.u.) for MSM-EE, –888.48761 (a.u.) for ONIOM-ME, and –888.49265 (a.u.) for MSM-ME.

Table 2 compares the computed and experimental free energies and enthalpy barriers. The ONIOM-EE and MSM-EE enthalpy barriers agree well with the experimental values, while the free energy barriers in both methods are slightly underestimated by a few kcal mol$^{-1}$. Previous work[56] has shown that the zero-point and thermal energies has the small contributions and reduce the enthalpy barrier by about 1.6 kcal mol–1, which is consistent with our work. However, ME underestimates the free energy and enthalpy barrier heights by about 7–8 kcal mol$^{-1}$ in this enzymatic reaction. Thus, EE treatment is necessary to obtain reliable results in this case.

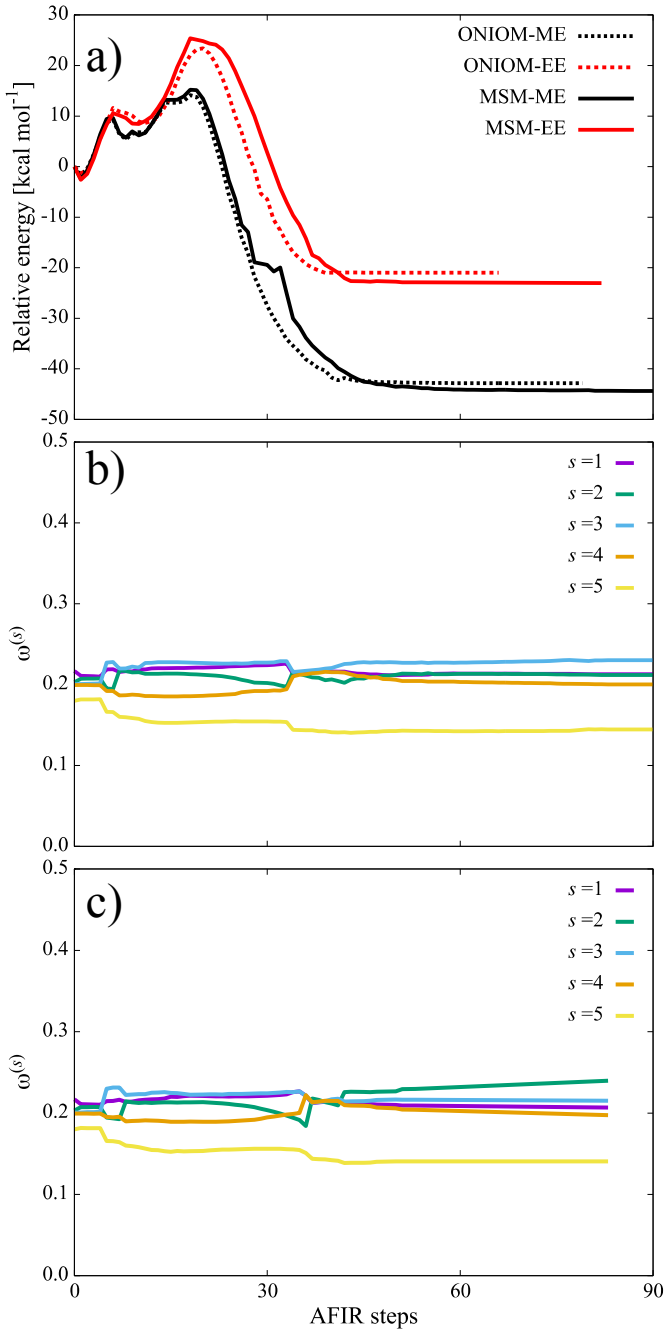


**Figure 4**. (a) Relative energies and variations of $\omega^{(s)}$ (b) in MSM-ME and (c) in MSM-EE along the AFIR path of Claisen rearrangement in an enzyme.

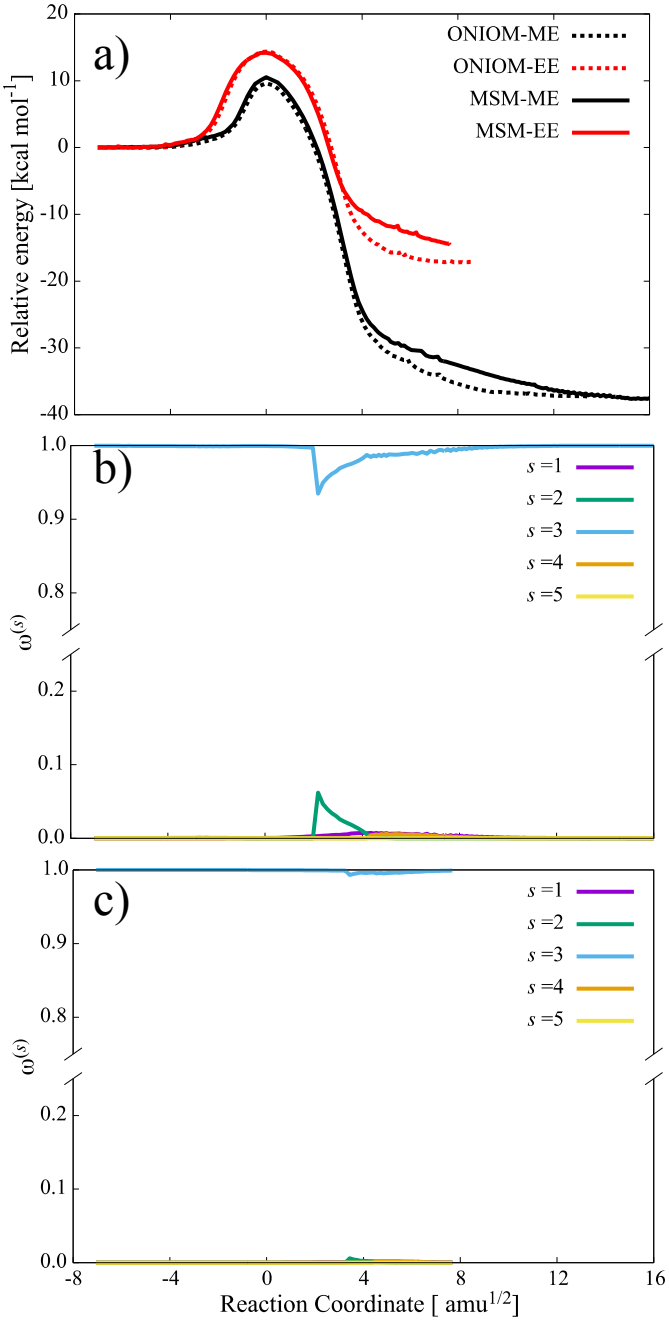


**Figure 5**. (a) Relative energies and variations of $\omega^{(s)}$ (b) in MSM-ME and (c) in MSM-EE along the IRC path of Claisen rearrangement in an enzyme.

**Table 2**. Potential energy (*E*), free energy (*G*), and enthalpy (*H*) barrier heights, and computed and experimental reaction energies of Claisen rearrangement in an enzyme. The units are kcal mol$^{-1}$.

| | Barrier height | | | Reaction energy | | |
|---|---|---|---|---|---|---|
| | *E* | *G* | *H* | *E* | *G* | *H* |
| ONIOM-ME | 9.6 | 8.9 | 8.5 | -37.3 | -37.1 | -36.9 |
| ONIOM-EE | 14.4 | 13.0 | 12.9 | -17.1 | -16.9 | -16.8 |
| MSM-ME | 10.5 | 7.1 | 9.6 | -37.7 | -37.2 | -37.2 |
| MSM-EE | 14.2 | 12.4 | 12.3 | -14.6 | -14.6 | -14.4 |
| Exp[63] | | 15.4 | 12.7 | | | |

### 4.3 Hydroxylation by *p*-hydroxybenzoate hydroxylase

Figures 6(a)–(c) show the relative energies from all methods and $\omega^{(s)}$ from MSM methods along the IRC path, which were obtained after AFIR path and LUP path optimizations as shown in Figures S6 and S7. The optimized geometric parameters at the stationary points are also shown in S8. The computed energy barriers and reaction energies are listed in Table 3. In Figure 6(a), the MSM-EE barrier is the highest. The order of obtained barrier heights is MSM-ME (7.7 kcal mol$^{-1}$) ≈ ONIOM-ME (8.4 kcal mol$^{-1}$) < ONIOM-EE (11.5 kcal mol$^{-1}$) < MSM-EE (13.7 kcal mol$^{-1}$). The enthalpy barriers computed in the ME schemes were underestimations, while those in the EE schemes agree well with experimental values. The difference between ONIOM-EE and MSM-EE along the IRC paths is that the most favorable surrounding structures are different. After AFIR and LUP path optimizations, the most favorable surrounding structure in MSM-EE is *s*=2 along the IRC path, while that of ONIOM-EE is *s*=1. This difference in most favorable surrounding structure lowers the total energy throughout the IRC path compared with that in ONIOM-EE. In this case, the total energies of the reactant states were –1574.39691 (a.u.) for ONIOM-EE and –1574.40567 (a.u.) for MSM-EE. The same tendency was observed in the ME schemes, where the total energies of the reactant states were –1574.42229 (a.u.) for ONIOM-ME and –1574.43430 (a.u.) for MSM-ME.

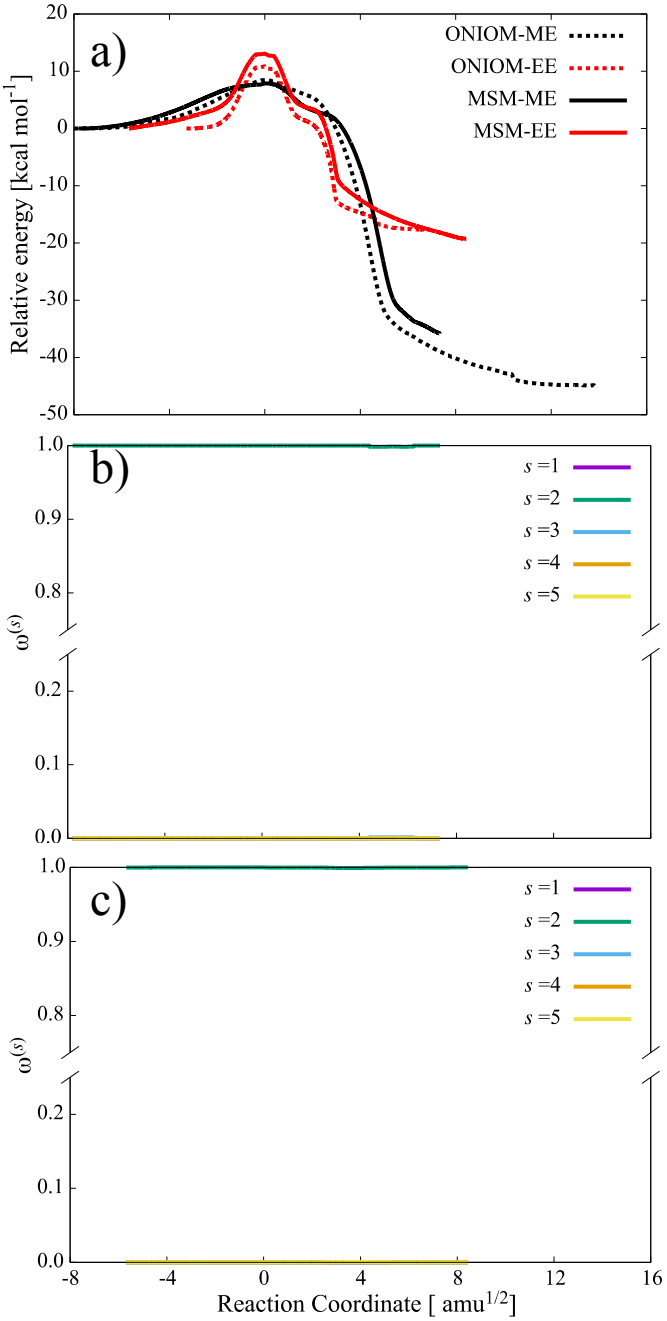


**Figure 6**. (a) Relative energies and variations of $\omega^{(s)}$ (b) in MSM-ME and c) in MSM-EE along the IRC path of hydroxylation in an enzyme.

**Table 3**. Potential energy (*E*), free energy (*G*), and enthalpy (*H*) barrier heights, and computed and experimental reaction energies of hydroxylation in an enzyme. The units are kcal $mol^{-1}$.

| | Barrier height | | | Reaction energy | | |
|---|---|---|---|---|---|---|
| | *E* | *G* | *H* | *E* | *G* | *H* |
| ONIOM-ME | 8.4 | 6.0 | 8.3 | -44.9 | -44.3 | -44.1 |
| ONIOM-EE | 11.5 | 10.1 | 10.6 | -17.6 | -18.9 | -18.3 |
| MSM-ME | 7.7 | 6.9 | 7.1 | -35.4 | -35.6 | -34.5 |
| MSM-EE | 13.7 | 13.2 | 12.8 | -19.3 | -19.0 | -19.4 |
| Exp[65] | | | 12.0 | | | |

### 4.4 Performances

Table 4 lists the numbers of QM force and Hessian calculations. The number of force calculations in ONIOM-ME is the smallest among the methods in this study. The computational cost is higher by roughly a factor of 1.1 for ONIOM-EE, 1.2 for MSM-ME, and 1.3 for MSM-EE compared with ONIOM-ME. The extra computational cost in MSM-EE is due to changes in the most favorable surrounding structures during the LUP path optimization. Nevertheless, an extra cost of only ~30% is acceptable in many practical applications considering its relatively high accuracy. MSM-EE can be a more reliable alternative to ONIOM-ME, ONIOM-EE, and MSM-ME and can be used routinely in geometrical optimization and reaction path calculation.

However, it should be noted about how we choose the surrounding structure. In this study, we studied with a single step reaction, and the surrounding structures were prepared for the reactant structure by replica-exchange MD. Discussions on how to generate, where to generate along the reaction path, and how to select the surrounding structures generated that gives more reliable results are necessary to be examined in the future. Tackling other technical subjects, such as the screened charge treatment for a short-range electrostatic interaction[78] and the polarizable embedding for taking

account of mutual polarization between the MM and QM parts,[33-37] will also extend the MSM framework further.

**Table 4**. Number of QM force and Hessian calculations required for AFIR, LUP, and IRC path optimizations in the various methods.

| | | AFIR | | LUP | | IRC | | Total | |
|---|---|---|---|---|---|---|---|---|---|
| | | Force | Hessian | Force | Hessian | Force | Hessian | Force | Hessian |
| Claisen in Aqueous solution | ONIOM-ME | 67 | 0 | 591 | 0 | 256 | 3 | 914 | 3 |
| | ONIOM-EE | 42 | 0 | 728 | 0 | 240 | 3 | 1010 | 3 |
| | MSM-ME | 74 | 0 | 688 | 0 | 257 | 3 | 1019 | 3 |
| | MSM-EE | 63 | 0 | 824 | 0 | 278 | 3 | 1165 | 3 |
| Claisen in Enzyme | ONIOM-ME | 79 | 0 | 273 | 0 | 235 | 3 | 587 | 3 |
| | ONIOM-EE | 66 | 0 | 391 | 0 | 256 | 3 | 713 | 3 |
| | MSM-ME | 92 | 0 | 400 | 0 | 231 | 3 | 723 | 3 |
| | MSM-EE | 82 | 0 | 399 | 0 | 259 | 3 | 740 | 3 |
| Hydroxylation in Enzyme | ONIOM-ME | 142 | 0 | 821 | 0 | 260 | 3 | 1223 | 3 |
| | ONIOM-EE | 106 | 0 | 995 | 0 | 266 | 3 | 1367 | 3 |
| | MSM-ME | 147 | 0 | 998 | 0 | 247 | 3 | 1392 | 3 |
| | MSM-EE | 95 | 0 | 1189 | 0 | 279 | 3 | 1563 | 3 |

## 5. Conclusions

We have proposed the MSM-EE method, which extends MSM-ME to electrostatic embedding, to efficiently perform the reaction path search in large molecular systems while taking account of both electron density polarization in the reaction center and the surrounding structural transition. To avoid performing the same number of QM calculations as for the surrounding structures in each optimization step, we included the partial charge scaled by the weights of multiple surrounding structures into the electronic Hamiltonian and took into account the polarization effect of the electron density in QM region. MSM-EE is a simple extension of ONIOM-EE; MSM-EE is expected to provide a lower energy path than ONIOM-EE when there is a finite multistructural

effect, while MSM-EE is also expected to provide a similar path to ONIOM-EE when there is little multistructural effect. Numerical tests showed that MSM-EE gave improved results compared with the ONIOM-ME, ONIOM-EE, and MSM-ME methods, and agreed well with experimental results. The additional computational cost is only increased by a factor of 1.3 compared with the least expensive ONIOM-ME. Therefore, it is expected that MSM-EE can be a more accurate alternative of ONIOM-ME and be applied to various chemical transformations routinely. Systematically uncovering cases in which the multistractural effects would be an exciting subject to be tackled in the future with MSM-EE. Another perspective is to use MSM-EE in an automated reaction path exploration and discuss kinetics of an enzyme reaction on the basis of a reaction path network consisting of numerous reaction paths.

**Corresponding Authors:** ki_suzuki@eis.hokudai.ac.jp (KS), smaeda@eis.hokudai.ac.jp (SM).

**Acknowledgments:** This work was supported by the Japan Science and Technology Agency with a Core Research for Evolutional Science and Technology (CREST) grant in the area of "Establishment of Molecular Technology towards the Creation of New Functions" at Hokkaido University (No. JPMJCR14L5) and a Grant-in-Aid for JST-ERATO (No. JPMJER1903). This work was supported by JSPS KAKENHI Grant Numbers 21K04974 (to K.S.) All calculations were performed at the Research Center Computational Science (RCCS), Okazaki, Japan. Mark Kurban from Edanz Group (https://www.jp.edanz.com/ac) edited a draft of this manuscript.

## References

(1) Houk, K. N; Cheong, P.H-Y. Computational prediction of small-molecule catalysts. *Nature,* **2008**, 455, 309–313. https://doi.org/10.1038/nature07368

(2) Thiel, W. Computational catalysis–past, present, and future. *Angew. Chem. Int. Ed*. 2014, 53, 8605–8613. https://doi.org/10.1002/anie.201402118

(3) Sameera, WMC; Maeda, S; Morokuma, K. Computational catalysis using the artificial force induced reaction method. *Acc Chem Res*. 2016, 49 (4) , 763–773. https://doi.org/10.1021/acs.accounts.6b00023

(4) A. L. Dewyer, A. J. Argüelles, P. M. Zimmerman, Methods for exploring reaction space in molecular systems. *WIREs Comput. Mol. Sci.* 8, e1354 (2018). doi:10.1002/wcms.1354

(5) G. N. Simm, A. C. Vaucher, M. Reiher, Exploration of reaction pathways and chemical transformation networks. *J. Phys. Chem*. A. 123, 385–399 (2019). doi:10.1021/acs.jpca.8b10007

(6) Maeda, S; Harabuchi, Y. Exploring paths of chemical transformations in molecular and periodic systems: An approach utilizing force. *WIREs Comput*. *Mol. Sci*. 2021, e1538. https://doi.org/10.1002 /wcms.1538

(7) Warshel, A.; Karplus, M. Calculation of Ground and Excited State Potential Surfaces of Conjugated Molecules.1 I. Formulation and Parametrization. *J. Am. Chem. Soc.* **1972**, *94* (16), 5612–5625. https://doi.org/10.1021/ja00771a014.

(8) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103* (2), 227–249. https://doi.org/10.1016/0022-2836(76)90311-9.

(9) Singh, U. C.; Kollman, P. A. A Combined ab Initio Quantum Mechanical and Molecular Mechanical Method for Carrying out Simulations on Complex Molecular Systems: Applications to the CH3Cl + Cl? Exchange Reaction and Gas Phase Protonation of Polyethers. *J. Comput. Chem.* **1986**, *7* (6), 718–730. https://doi.org/10.1002/jcc.540070604.

(10) Field, M. J.; Bash, P. A.; Karplus, M. A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations. *J. Comput. Chem.* **1990**, *11* (6), 700–733. https://doi.org/10.1002/jcc.540110605.

(11) Maseras, F.; Morokuma, K. IMOMM: A New Integrated ab Initio + Molecular Mechanics Geometry Optimization Scheme of Equilibrium Structures and Transition States. *J. Comput. Chem.* **1995**, *16* (9), 1170–1179. https://doi.org/10.1002/jcc.540160911.

(12) Ryde, U. The Coordination of the Catalytic Zinc Ion in Alcohol Dehydrogenase Studied by Combined Quantum-Chemical and Molecular Mechanics Calculations. *J. Comput. Aided. Mol. Des.* **1996**, *10* (2), 153–164. https://doi.org/10.1007/BF00402823.

(13) Kästner, J.; Thiel, S.; Senn, H. M.; Sherwood, P.; Thiel, W. Exploiting QM/MM Capabilities in Geometry Optimization: A Microiterative Approach Using Electrostatic Embedding. *J. Chem. Theory Comput.* **2007**, *3* (3), 1064–1072. https://doi.org/10.1021/ct600346p.

(14) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem. Int. Ed.* **2009**, *48* (7), 1198–1229. https://doi.org/10.1002/anie.200802019.

(15) Chung, L. W.; Sameera, W. M. C.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z.; Liu, F.; et al. The ONIOM Method and Its Applications. *Chem. Rev.* **2015**, *115* (12), 5678–5796. https://doi.org/10.1021/cr5004419.

(16) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.;Farkas,O.; Foresman,J.B.; Ortiz,J.V.;Cioslowski,J.;Fox,D.J. Gaussian 09, revision D.01; Gaussian, Inc.: Wallingford, CT, 2009.

(17) Isborn, C. M.; Götz, A. W.; Clark, M. A.; Walker, R. C.; Martínez, T. J. Electronic Absorption Spectra from MM and Ab Initio QM/MM Molecular Dynamics: Environmental Effects on the Absorption Spectrum of Photoactive Yellow Protein. *J. Chem. Theory Comput.* **2012**, *8* (12), 5092–5106. https://doi.org/10.1021/ct3006826.

(18) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; et al. Advances in Molecular Quantum Chemistry Contained in the Q-Chem 4 Program Package. *Mol. Phys.* **2015**, *113* (2), 184–215. https://doi.org/10.1080/00268976.2014.952696.

(19) Salahub, D. R.; Noskov, S. Y.; Lev, B.; Zhang, R.; Ngo, V.; Goursot, A.; Calaminici, P.; Köster, A. M.; Alvarez-Ibarra, A.; Mejía-Rodríguez, D.; et al. QM/MM Calculations with DeMon2k. *Molecules* **2015**, *20* (3), 4780–4812. https://doi.org/10.3390/molecules20034780.

(20) Takano, Y.; Nakata, K.; Yonezawa, Y.; Nakamura, H. Development of Massive Multilevel Molecular Dynamics Simulation Program, Platypus (PLATform for DYnamic Protein Unified Simulation), for the Elucidation of Protein Functions. *J. Comput. Chem.* **2016**, *37* (12), 1125–1132. https://doi.org/10.1002/jcc.24318.

(21) Neese, F. Software Update: The ORCA Program System, Version 4.0. *WIREs Comput. Mol. Sci.* **2018**, *8* (1) e1327. https://doi.org/10.1002/wcms.1327.

(22) Shoemaker, J. R.; Gordon, M. S. SIMOMM: An Integrated Molecular Orbital/Molecular Mechanics Optimization Scheme for Surfaces. *J. Phys. Chem.* A **1999**, *103* (17), 3245–3251. https://doi.org/10.1021/jp982600e.

(23) Riahi, S.; Rowley, C. N. The CHARMM-TURBOMOLE Interface for Efficient and Accurate QM/MM Molecular Dynamics, Free Energies, and Excited State Properties. *J. Comput. Chem.* **2014**, *35* (28), 2076–2086. https://doi.org/10.1002/jcc.23716.

(24) Ootani, Y.; Akinaga, Y.; Nakajima, T. Theoretical Investigation of Enantioselectivity of Cage-like Supramolecular Assembly: The Insights into the Shape Complementarity and Host-Guest Interaction. *J. Comput. Chem.* **2015**, *36* (7), 459–466. https://doi.org/10.1002/jcc.23821.

(25) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26* (16), 1668–1688. https://doi.org/10.1002/jcc.20290.

(26) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26* (16), 1701–1718. https://doi.org/10.1002/jcc.20291.

(27) Walker, R. C.; Crowley, M. F.; Case, D. A. The Implementation of a Fast and Accurate QM/MM Potential Method in Amber. *J. Comput. Chem.* **2008**, *29* (7), 1019–1031.

https://doi.org/10.1002/jcc.20857.

(28) Metz, S.; Kästner, J.; Sokol, A. A.; Keal, T. W.; Sherwood, P. ChemShell-a Modular Software Package for QM/MM Simulations. *WIREs Comput. Mol. Sci.* **2014**, *4* (2), 101–110. https://doi.org/10.1002/wcms.1163.

(29) Vilseck, J. Z.; Kostal, J.; Tirado-Rives, J.; Jorgensen, W. L. Application of a BOSS - Gaussian Interface for QM/MM Simulations of Henry and Methyl Transfer Reactions. *J. Comput. Chem.* **2015**, *36* (27), 2064–2074. https://doi.org/10.1002/jcc.24045.

(30) Shiga, M. PIMD, Version 2.1. http://ccse.Jaea.go.jp/ja/download/pimd/index.en.html (2016).

(31) Melo, M. C. R.; Bernardi, R. C.; Rudack, T.; Scheurer, M.; Riplinger, C.; Phillips, J. C.; Maia, J. D. C.; Rocha, G. B.; Ribeiro, J. V; Stone, J. E.; et al. NAMD Goes Quantum: An Integrative Suite for Hybrid Simulations. *Nat. Methods* **2018**, *15* (5), 351–354. https://doi.org/10.1038/nmeth.4638.

(32) Yagi, K.; Yamada, K.; Kobayashi, C.; Sugita, Y. Anharmonic Vibrational Analysis of Biomolecules and Solvated Molecules Using Hybrid QM/MM Computations. *J. Chem. Theory Comput.* **2019**, *15* (3), 1924–1938. https://doi.org/10.1021/acs.jctc.8b01193.

(33) Thompson, M. A.; Schenter, G. K. Excited States of the Bacteriochlorophyll b Dimer of Rhodopseudomonas Viridis: A QM/MM Study of the Photosynthetic Reaction Center That Includes MM Polarization. *J. Phys. Chem.* **1995**, *99* (17), 6374–6386. https://doi.org/10.1021/j100017a017.

(34) Bakowies, D.; Thiel, W. Semiempirical Treatment of Electrostatic Potentials and Partial Charges in Combined Quantum Mechanical and Molecular Mechanical Approaches. *J. Comput. Chem.* **1996**, *17* (1), 87–108. https://doi.org/10.1002/(SICI)1096-987X(19960115)17:1<87::AID-JCC8>3.0.CO;2-X.

(35) Giovannini, T.; Puglisi, A.; Ambrosetti, M.; Cappelli, C. Polarizable QM/MM Approach with Fluctuating Charges and Fluctuating Dipoles: The QM/FQFμ Model. J. Chem. Theory Comput. 2019, 15, 2233-2245.https://doi.org/10.1021/acs.jctc.8b01149.

(36) Lemkul, J. A.; Huang, J.; Roux, B.; MacKerell, A. D. An Empirical Polarizable Force Field Based on the Classical Drude Oscillator Model: Development History and Recent Applications. Chem. Rev., 2016, 116, 4983-5013. https://doi.org/10.1021/acs.chemrev.5b00505

(37) Wu, X.; Teuler, J-M.; Cailliez, F.; Clavaguéra, C.; Salahub, D. R.; de la Lande, A.Simulating Electron Dynamics in Polarizable Environments. Journal of Chemical Theory and Computation 2017, 13, 3985-4002. https://doi.org/10.1021/acs.jctc.7b00251

(38) Pezeshki, S.; Lin, H. Recent developments in QM/MM methods towards open-boundary multi-scale simulations. Mol. Sim., 2014, 41, 168-189. https://doi.org/10.1080/08927022.2014.911870

(39) Vreven, T.; Morokuma, K. Investigation of the S0→S1 Excitation in Bacteriorhodopsin with the ONIOM(MO:MM) Hybrid Method. *Theor. Chem. Acc.* **2003**, *109* (3), 125–132. https://doi.org/10.1007/s00214-002-0418-y.

(40) Vreven, T.; Byun, K. S.; Komáromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. Combining Quantum Mechanics Methods with Molecular Mechanics Methods in ONIOM. *J. Chem. Theory Comput.* **2006**, *2* (3), 815–826. https://doi.org/10.1021/ct050289g.

(41) Lin, H.; Truhlar, D. G. QM/MM: What Have We Learned, Where Are We, and Where Do We Go from Here? *Theor. Chem. Acc.* **2007**, *117* (2), 185–199. https://doi.org/10.1007/s00214-006-0143-z.

(42) Lundberg, M.; Borowski, T. Oxoferryl Species in Mononuclear Non-Heme Iron Enzymes: Biosynthesis, Properties and Reactivity from a Theoretical Perspective. *Coord. Chem. Rev.* **2013**, *257* (1), 277–289. https://doi.org/10.1016/j.ccr.2012.03.047.

(43) Vreven, T.; Morokuma, K.; Farkas, Ö.; Schlegel, H. B.; Frisch, M. J. Geometry Optimization with QM/MM, ONIOM, and Other Combined Methods. I. Microiterations and Constraints. *J. Comput. Chem.* **2003**, *24* (6), 760–769. https://doi.org/10.1002/jcc.10156.

(44) Polyak, I.; Boulanger, E.; Sen, K.; Thiel, W. A Microiterative Intrinsic Reaction Coordinate Method for Large QM/MM Systems. *Phys. Chem. Chem. Phys.* **2013**, *15* (34), 14188. https://doi.org/10.1039/c3cp51669e.

(45) Maeda, S.; Abe, E.; Hatanaka, M.; Taketsugu, T.; Morokuma, K. Exploring Potential Energy Surfaces of Large Systems with Artificial Force Induced Reaction Method in Combination with ONIOM and Microiteration. *J. Chem. Theory Comput.* **2012**, *8* (12), 5058–5063. https://doi.org/10.1021/ct300633e.

(46) Maeda, S.; Ohno, K.; Morokuma, K. An Automated and Systematic Transition Structure Explorer in Large Flexible Molecular Systems Based on Combined Global Reaction Route

Mapping and Microiteration Methods. *J. Chem. Theory Comput.* **2009**, *5* (10), 2734–2743. https://doi.org/10.1021/ct9003383.

(47) Yasuda, R.; Noji, H.; Yoshida, M.; Kinosita, K.; Itoh, H. Resolution of Distinct Rotational Substeps by Submillisecond Kinetic Analysis of F1-ATPase. *Nature* **2001**, *410* (6831), 898–904. https://doi.org/10.1038/35073513.

(48) Kędzierski, P.; Moreton, K.; Clarke, A. R.; Holbrook, J. J. The A245K Mutation Exposes Another Stage of the Bacterial L-Lactate Dehydrogenase Reaction Mechanism. *Biochemistry* **2001**, *40* (24), 7247–7252. https://doi.org/10.1021/bi0026775.

(49) Boehr, D. D.; McElheny, D.; Dyson, H. J.; Wrightt, P. E. The Dynamic Energy Landscape of Dihydrofolate Reductase Catalysis. *Science* **2006**, *313* (5793), 1638–1642. https://doi.org/10.1126/science.1130258.

(50) Henzler-Wildman, K. A.; Thai, V.; Lei, M.; Ott, M.; Wolf-Watz, M.; Fenn, T.; Pozharski, E.; Wilson, M. A.; Petsko, G. A.; Karplus, M.; et al. Intrinsic Motions along an Enzymatic Reaction Trajectory. *Nature* **2007**, *450* (7171), 838–844. https://doi.org/10.1038/nature06410

(51) Suzuki, K.; Morokuma, K.; Maeda, S. Multistructural Microiteration Technique for Geometry Optimization and Reaction Path Calculation in Large Systems. *J. Comput. Chem.* **2017**, *38* (26), 2213–2221. https://doi.org/10.1002/jcc.24857.

(52) Suzuki, K.; Maeda, S.; Morokuma, K. Roles of Closed- and Open-Loop Conformations in Large-Scale Structural Transitions of l -Lactate Dehydrogenase. *ACS Omega* **2019**, *4* (1), 1178–1184. https://doi.org/10.1021/acsomega.8b02813.

(53) Lever, G.; Cole, D. J.; Lonsdale, R.; Ranaghan, K. E.; Wales, D. J.; Mulholland, A. J.; Skylaris, C. K.; Payne, M. C. Large-Scale Density Functional Theory Transition State Searching in Enzymes. *J. Phys. Chem. Lett.* **2014**, *5* (21), 3614–3619. https://doi.org/10.1021/jz5018703.

(54) Crespo, A.; Scherlis, D. A.; Martí, M. A.; Ordejón, P.; Roitberg, A. E.; Estrin, D. A. A DFT-Based QM-MM Approach Designed for the Treatment of Large Molecular Systems: Application to Chorismate Mutase. *J. Phys. Chem. B* **2003**, *107* (49), 13728–13736. https://doi.org/10.1021/jp036236h.

(55) Claeyssens, F.; Ranaghan, K. E.; Lawan, N.; MacRae, S. J.; Manby, F. R.; Harvey, J. N.; Mulholland, A. J. Analysis of Chorismate Mutase Catalysis by QM/MM Modelling of

Enzyme-Catalysed and Uncatalysed Reactions. *Org. Biomol. Chem.* **2011**, *9* (5), 1578–1590. https://doi.org/10.1039/c0ob00691b.

(56) Claeyssens, F.; Harvey, J. N.; Manby, F. R.; Mata, R. A.; Mulholland, A. J.; Ranaghan, K. E.; Schütz, M.; Thiel, S.; Thiel, W.; Werner, H. J. High-Accuracy Computation of Reaction Barriers in Enzymes. *Angew. Chem. Int. Ed.* **2006**, *45* (41), 6856–6859. https://doi.org/10.1002/anie.200602711.

(57) Ishida, T. Effects of Point Mutation on Enzymatic Activity: Correlation between Protein Electronic Structure and Motion in Chorismate Mutase Reaction. *J. Am. Chem. Soc.* **2010**, *132* (20), 7104–7118. https://doi.org/10.1021/ja100744h.

(58) Carlson, H. A.; Jorgensen, W. L. Monte Carlo Investigations of Solvent Effects on the Chorismate to Prephenate Rearrangement. *J. Am. Chem. Soc.* **1996**, *118* (35), 8475–8484. https://doi.org/10.1021/ja961500o.

(59) Bistni, G; Polyak, I.; Sparta, M.; Thiel, W.; Neese, F. Toward Accurate QM/MM Reaction Barriers with Large QM Regions Using Domain Based Pair Natural Orbital Coupled Cluster Theory. *J. Chem. Thory Comput.* **2018**, 14, 3524-3531. https://doi.org/10.1021/acs.jctc.8b00348

(60) Andrews, P. R.; Smith, G. D.; Young, I. G. Transition-State Stabilization and Enzymic Catalysis. Kinetic and Molecular Orbital Studies of the Rearrangement of Chorismate to Prephenate. *Biochemistry* **1973**, *12* (18), 3492–3498. https://doi.org/10.1021/bi00742a022.

(61) Copley, S. D.; Knowles, J. R. The Conformational Equilibrium of Chorismate in Solution: Implications for the Mechanism of the Non-Enzymic and the Enzyme-Catalyzed Rearrangement of Chorismate to Prephenate. *J. Am. Chem. Soc.* **1987**, *109* (16), 5008–5013. https://doi.org/10.1021/ja00250a040.

(62) Chook, Y. M.; Ke, H.; Lipscomb, W. N. Crystal Structures of the Monofunctional Chorismate Mutase from Bacillus Subtilis and Its Complex with a Transition State Analog. *Proc. Natl. Acad. Sci.* **1993**, 90 (18), 8600–8603. https://doi.org/10.1073/pnas.90.18.8600.

(63) Kast, P.; Asif-Ullah, M.; Hilvert, D. Is Chorismate Mutase a Prototypic Entropy Trap? - Activation Parameters for the Bacillus Subtilis Enzyme. *Tetrahedron Lett.* **1996**, *37* (16), 2691–2694. https://doi.org/10.1016/0040-4039(96)00338-3.

(64) Wright, S. K.; Declue, M. S.; Mandal, A.; Lee, L.; Wiest, O.; Cleland, W. W.; Hilvert, D. Isotope Effects on the Enzymatic and Nonenzymatic Reactions of Chorismate. *J. Am. Chem. Soc.* **2005**, *127* (37), 12957–12964. https://doi.org/10.1021/ja052929v.

(65) Van Berekel, W. J. H; Müller, F. The Temperature and pH Dependence of Some Properties of p-Hydroxybenzoate Hydroxylase from Pseudomonas Fluorescens. *Eur. J. Biochem*. **1989**, 179(2), 307-314. https://doi.org/10.1111/j.1432-1033.1989.tb14556.x

(66) Ortiz-Maldonado, M.; Cole, L. J.; Dumas, S. M.; Entsch, B.; Ballou, D. P. Increased Positive Electrostatic Potential in p-Hydroxybenzoate Hydroxylase Accelerates Hydroxylation but Slows Turnover. *Biochemistry* **2004**, 43, 1569-1579. https://doi.org/10.1021/bi030193d

(67) Schreuder H. A.; Prick, P. A. J.; Wierenga, R.K.; Vriend, G.; Wilson, K. S.; Hol, W. G. J.; Drenth, J.; Crystal Structure of the p-Hydroxybenzoate Hydroxylase-Substrate Complex Refined at 1.9 Å Resolution: Analysis of the Enzyme-Substrate and Enzyme-Product Complexes. *J. Mol. Biol.* 1989, 208(4), 679-696. https://doi.org/10.1016/0022-2836(89)90158-7

(68) Li, H.; Robertson, A. D.; Jensen, J. H. Very Fast Empirical Prediction and Rationalization of Protein PKa Values. *Proteins Struct. Funct. Bioinforma.* **2005**, *61* (4), 704–721. https://doi.org/10.1002/prot.20660.

(69) Bas, D. C.; Rogers, D. M.; Jensen, J. H. Very Fast Prediction and Rationalization of PKa Values for Protein-Ligand Complexes. *Proteins Struct. Funct. Bioinforma.* **2008**, *73* (3), 765–783. https://doi.org/10.1002/prot.22102.

(70) Maeda, S.; Harabuchi, Y.; Sumiya, Y.; Takagi, M.; Suzuki, K.; Hatanaka, M.; Osada, Y.; Taketsugu, T.; Morokuma, K.; Ohno, K. GRRM, a developmental version, see http://iqce.jp/index_e.shtml (accessed date June 19, 2018).

(71) Maeda, S.; Harabuchi, Y.; Takagi, M.; Saita, K.; Suzuki, K.; Ichino, T.; Sumiya, Y.; Sugiyama, K.; Ono, Y. Implementation and Performance of the Artificial Force Induced Reaction Method in the GRRM17 Program. *J. Comput. Chem*. 2018, 39 (4), 233–251. https://doi.org/10.1002/jcc.25106.

(72) Maeda, S.; Morokuma, K. Communications: A Systematic Method for Locating Transition Structures of A+B→X Type Reactions. *J. Chem. Phys.* 2010, 132 (24), 241102. https://doi.org/10.1063/1.3457903.

(73) Choi, C.; Elber, R. Reaction Path Study of Helix Formation in Tetrapeptides: Effect of Side Chains. *J. Chem. Phys.* **1991**, *94* (1), 751–760. https://doi.org/10.1063/1.460343.

(74) Izsák, R; Auer, A. A.; Neese, F. An overlap fitted chain of spheres exchange method. *J. Chem. Phys.* **2011**, 135 (14), 144105. https://doi.org/10.1063/1.3646921.

(75) Stoychev, G. L.; Auer, A. A.; Neese, F. Automatic Generation of Auxiliary Basis Sets. *J. Chem. Theory Comput.* **2017**, 13 (2), 554–562. https://doi.org/10.1021/acs.jctc.6b01041.

(76) Thellamurege, N, M.; Hirao, H. Effect of Protein Environment within Cytochrome P450cam Evaluated Using a Polarizable-Embedding QM/MM method. J. Phys. Chem. B, 2014, 118 (8), 2084-2092. https://doi.org/10.1021/jp412538n

(77) Seminario, J. M. Calculation of Intramolecular Force Fields from Second-Derivative Tensors. *Int. J. Quantum Chem*. **1996**, 60 (7), 1271–1277. https://doi.org/10.1002/(SICI)1097-461X(1996)60:7<1271::AID-QUA8>3.0.CO;2-W.

(78) Rackers, J, A.; Wang, Q.; Liu, C.; Piquemal, J.-P.; Ren, P.; Ponder, J. W.;. An Optimized Charge Penetration Model for Use with the AMOEBA Force Field. *Phys, Chem. Chem. Phys.* **2017**, 19 (1), 276–291. https://doi.org/10.1039/C6CP06017J.